Proc. Natl. Acad. Sci. USA
Vol. 83, pp. 5459–5463, August 1986
Biochemistry

yes-related protooncogene, *syn*, belongs to the protein-tyrosine kinase family

(kinase family/*yes* gene/cDNA cloning)

KENTARO SEMBA*, MAKOTO NISHIZAWA*, NOBUYUKI MIYAJIMA*, MICHIHIRO C. YOSHIDA†, JUN SUKEGAWA*, YUJI YAMANASHI*, MOTOMICHI SASAKI†, TADASHI YAMAMOTO*, AND KUMAO TOYOSHIMA*

*The Institute of Medical Science, The University of Tokyo 4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan; and †Chromosome Research Unit, Hokkaido University, Kita-ku, Sapporo 060, Japan

Communicated by Peter K. Vogt, April 14, 1986

ABSTRACT **An *src/yes*-related novel gene named *syn* (*SYN* in human gene nomenclature) has been identified in the human genome on chromosome 6 and characterized by molecular cloning. Nucleotide sequence analysis of cDNA clones showed that the c-*syn* gene could encode a protein-tyrosine kinase that is very similar in primary structure to the v-yes and human c-src proteins. A 2.8-kilobase transcript of the c-*syn* gene, which differs in size from those of the c-*yes*, c-*src*, and c-*fgr* genes, was observed in various cell types. These results show that *syn* is a new member of the tyrosine kinase oncogene family.**

At least eight retroviral oncogenes (*src, yes, fgr, fps/fes, abl, ros, fms*, and *erbB*) have been classified as members of a "tyrosine kinase family" by the fact that their products exhibit protein-tyrosine kinase activity (1). This activity is also associated with receptors for growth factors such as epidermal growth factor (2), insulin (3), and platelet-derived growth factor (4). Recent findings showed that there is overlap between these two types of protein-tyrosine kinases. For example, c-*erbB* (*ERBB*)‡ (5) and c-*fms* (*FMS*)‡ (6) genes encode the epidermal growth factor receptor and mononuclear phagocyte growth factor (CSF-1) receptor, respectively. Thus, at least several protooncogenes of the tyrosine kinase family are involved in the regulation of normal cell growth, and subversion of the protooncogene expression by retroviral induction or possibly by mutation mediates abnormal cell growth leading to neoplastic transformation. However, the function of "nonreceptor"-type protooncogenes of the tyrosine kinase family, such as the c-*src* (*SRC*)‡ gene is unknown. We previously cloned the human c-*yes-1* (*YES1*)‡, c-*yes-2* (*YES2*)‡ (pseudogene of c-*yes-1*), and c-*fgr*‡ genes, which are localized on the human chromosomes 18q21.3 (7, 8), 6 (7), and 1p36.1–36.2 (9, 10), respectively. Besides these genes, we identified an *src/yes*-related novel gene (*syn*)‡ in the human genome. Here we show the molecular cloning and characterization of the c-*syn* gene.

MATERIALS AND METHODS

Cells and Tissues. A431 cells, UCVA cells (11), and human embryo fibroblasts were maintained in Dulbecco's modified Eagle's medium with 10% fetal calf serum. K562 cells, IM9 cells, and FL18 cells (12) were maintained in RPMI 1640 medium with 10% fetal calf serum.

Isolation of Clones. A human genomic library from placental DNA and a cDNA library from human embryo fibroblasts were constructed by using phage λ Charon 4A and λgt10, respectively, as vectors (13, 14). Screening of the genomic library was as described (9). The cDNA library (3.7×10^5 plaques) was screened with a ^{32}P-labeled c-*syn*-specific probe derived from the c-*syn* genomic clone (nick-translated to a specific activity of 2×10^8 cpm/μg of DNA) under stringent conditions (15).

Nucleotide Sequence Analysis. The nucleotide sequence was determined by the dideoxy chain-termination method (16) with modification (17, 18) in conjunction with plasmid pUC19.

Blot-Hybridization Analysis of DNA and RNA. High molecular weight DNAs were digested with restriction endonucleases *Hin*dIII under the conditions recommended by suppliers (Takara Shuzo Co., Kyoto, Japan) and fractionated by electrophoresis in 0.9% agarose gels. The fragments were subjected to Southern blot hybridization with the c-*syn* probe or with a v-*myb* probe as described (15). Poly(A)$^+$ RNAs were electrophoresed in 1% agarose gel containing 2.2 M formaldehyde and were subjected to blot hybridization as described (15). The DNA probes were labeled with ^{32}P by nick-translation to a specific activity of 2×10^8 cpm/μg of DNA.

Karyotype Analysis of Human–Mouse Hybrid Cells. Somatic cell hybrids were generated by fusion (19) of human embryo fibroblasts with mouse FM3A cells. The human chromosome content was determined by the method of differential staining of human and mouse chromosomes (20).

RESULTS AND DISCUSSION

Identification of a *yes*-Related Gene. With a 1.5-kilobase-pair (kbp) v-*yes* DNA segment (7) as a screening probe, we isolated 26 independent clones from a human gene library (9). The restriction maps of 3 of them were distinct from those of the genomic clones representing the c-*yes-1*, c-*yes-2*, c-*fgr*, and c-*src* genes (refs. 9 and 21; unpublished data), which suggested that these 3 clones represent a novel v-*yes*-related gene. We named this the *syn* gene (*src/yes*-related novel protooncogene). By analyzing the nucleotide sequence of the cloned insert that hybridized with the v-*yes* probe, one putative exon flanked by the splicing consensus sequence (22) was identified (data not shown). The sequence of the exon (180 bp) is highly homologous with corresponding sequence of the v-*yes* gene (79% homology). Interestingly, the splicing sites of this putative exon are identical to those of the corresponding exons of the chicken (23) and human

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: kbp, kilobase pair(s); KD probe, *Kpn* I–*Dra* I DNA fragment from the c-*syn* genomic clone.

‡The name recently assigned to the human locus appears in parentheses; although c-fps, c-fgr, and c-syn have not yet been assigned human gene names, guidelines suggest that the c will be omitted and the rest will become italicized capital letters.

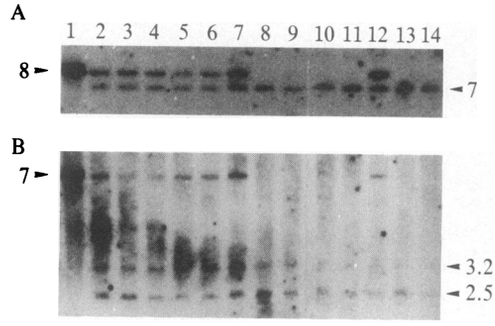


FIG. 1. Identification of human c-*syn* in DNAs of human–mouse somatic cell hybrids. (*A*) Nitrocellulose filters containing *Hin*dIII digests of the 12 hybrid DNAs (10 μg per lane) were probed with c-*syn*-specific KD DNA under stringent conditions. (*B*) After removal of the first probe, the filters were then hybridized with the v-*myb* probe under relaxed conditions (15). DNAs in lanes: 1, human placenta; 14, mouse FM3A cells; 2–13, human–mouse cell hybrids: 1a, II-5, III-1, A1, II-6, Bm, 1B1, 3-2, 7-2, 7D4, 6-3, and 3D3. ^{32}P-labeled fragments from *Hin*dIII digestion of phage λ DNA were used as a size marker. Sizes are shown in kbp.

(24) c-*src* (exon 8), human c-*fgr* (9), and human c-*yes-1* (unpublished data) genes. Thus, the c-*syn* gene is evolutionarily close to the c-*src*, c-*fgr*, and c-*yes* genes, with which it shows a common intron–exon organization that differs from those of other members of the c-*src* family such as c-*abl* (ABL)‡ (25), c-*fps*‡ (26), c-*raf* (*RAF*)/*mil*‡ (27), and c-*erbB-2* (*ERBB2*) (15).

Allocation of the c-*syn* Gene to Human Chromosome 6. The chromosomal location of the c-*syn* gene was determined by Southern blot analysis of 14 human–mouse hybrid cells with ^{32}P-labeled *Kpn* I–*Dra* I (KD) fragment generated from the c-*syn* genomic clone (Fig. 1*A* and Table 1). This probe contains the whole exon described above and hybridized with an 8-kbp *Hin*dIII fragment of the human c-*syn* gene (Fig. 1*A*). The probe also cross-hybridized with the 7-kbp fragment of the mouse c-*syn* gene. The data showed that the c-*syn* gene is located on human chromosome 6, while the c-*yes-1*, c-*src*, and c-*fgr* genes have been localized on 18q21.3 (7, 8), 20q12–13 (28), and 1p36.1–36.2 (9, 10), respectively. The chromosome localization of the c-*syn* gene was confirmed by probing the DNAs of hybrid cells with the v-*myb* DNA, whose human homolog (*MYB*)‡ has been assigned to chromosome 6. The v-*myb* probe reacted with a 7-kbp *Hin*dIII fragment of the human c-*myb* gene and with 3.2-kbp and 2.5-kbp fragments of the mouse c-*myb* gene (Fig. 1*B*). We previously mapped a v-*yes*-related pseudogene, c-*yes-2*, on chromosome 6 (7). However, the c-*syn* gene is distinct from the c-*yes-2* gene because it has different nucleotide sequences in corresponding regions. For example, the nucleotide sequence of the c-*yes-2* gene that corresponds to the sequence 1369–1398 of c-*syn* (Fig. 2*B*) is TGG GAA ATC CCT TGA TAA TCT TTG CGA CTA. A deletion of the distal half of chromosome 6q is often observed in acute lymphoblastic leukemia (29). Translocation between chromosomes 6 and 14 is associated with ovarian papillary adenocarcinomas (29). The precise location of the c-*syn* gene and further analysis of the c-*syn* locus in these tumors are required.

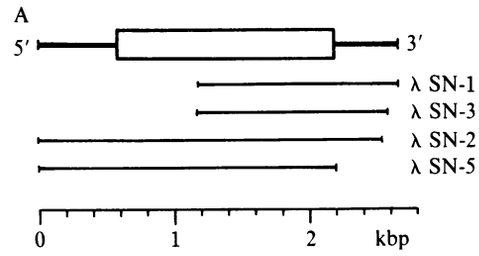
cDNA Cloning and Molecular Characterization of c-*syn*. RNA blot analysis with the ^{32}P-labeled KD fragment revealed a 2.8-kb c-syn mRNA in human embryo fibroblasts (see below). A cDNA library constructed from mRNA from human embryo fibroblasts was screened with the KD probe. Among four positive clones, λSN-2 had the longest insert of nearly full length (2.6 kbp), consisting of one *Eco*RI insert (Fig. 2*A*). The 2.6-kbp *Eco*RI insert was recloned (pSN-2) into plasmid vector pUC19 and then subjected to nucleotide sequence analysis. The longest open reading frame starting with a methionine codon at position 580 in the nucleotide sequence encodes a 537-amino acid polypeptide that includes the 59-amino acid residues of the putative exon described above (positions 289–347, Fig. 2*B*). Although the flanking nucleotides of the predicted initiation codon ATG do not show a perfect match with Kozak's consensus sequence (30), this codon seems to be the initiation codon because a stop codon (TAG) is found in the frame, immediately upstream in the 5′ untranslated region. The second ATG codon in the same open reading frame is located 583 nucleotides downstream. Translation of the nucleotide sequence between the first and second methionine codons gives a peptide with an amino acid sequence homologous to that of the corresponding portion of the v-yes protein (see below). The predicted coding sequence of 1611 bp is flanked by 5′- and 3′-untranslated sequences of 580 bp and 454 bp, respectively. The extreme 3′ sequence must have been lost from pSN-2 during the cloning procedure because we could not find any poly(A) sequence at the 3′-untranslated region. It is uncertain whether the AATAAA sequence present near the 3′ terminus is used as a poly(A) addition signal or not. A primary translation product of the c-*syn* gene was calculated to have a relative molecular mass of 60,761.

syn, Structurally Characterized as a New Member of the Tyrosine Kinase Family. The 537-amino acid sequence deduced from the cDNA clone pSN-2 specifies a protein closely related to, but distinct from, protein-tyrosine kinases encoded by the *src*, *yes*, and *fgr* genes. Fig. 3 shows a comparison

Table 1. Segregation of the c-*syn* gene with human chromosomes in human–mouse cell hybrids

	Chromosome																							
Hybrid	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X	Y	c-*syn**
1a	−	−	+	+	−	+	−	+	−	−	+	+	+	+	+	+	+	−	−	+	+	+	+	+	+
II-5	−	+	+	+	+	+	+	−	−	+	−	+	−	+	+	−	+	+	−	+	+	+	+	−	+
III-1	−	−	+	+	+	+	+	+	−	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+
A1	−	−	+	+	+	+	−	+	−	−	+	−	+	+	+	−	+	+	+	+	+	+	+	+	+
II-6	+	+	−	+	+	+	−	+	−	−	+	−	+	+	+	−	+	+	−	+	+	+	−	−	+
Bm	+	+	+	+	+	+	+	+	−	+	+	+	−	+	+	+	+	+	+	+	+	+	+	+	+
1B1	+	−	+	−	−	−	+	−	−	−	−	−	+	+	−	−	+	−	+	+	−	−	−	−	−
3-2	−	−	−	−	−	−	+	−	−	−	−	−	+	+	−	−	+	+	−	−	−	−	−	−	−
7-2	−	−	+	−	+	−	−	−	−	−	−	+	+	+	−	−	+	−	−	+	+	−	−	−	−
7D4	−	−	−	+	+	−	+	+	−	−	+	+	+	+	+	+	+	+	+	+	+	−	+	−	−
6-3	−	−	−	+	+	+	+	+	−	−	+	+	−	+	+	−	+	−	+	+	+	+	−	−	+
3D3	−	−	−	−	−	−	−	+	−	−	+	−	−	+	+	−	−	−	−	+	+	−	+	−	−
1-3	−	−	−	+	−	+	−	−	−	−	−	−	−	+	−	−	+	−	−	−	−	−	−	−	+
1-4	+	−	−	+	−	−	+	+	−	+	+	−	+	+	−	−	+	−	+	+	+	+	+	−	−

*Hybridization analysis is shown in Fig. 1, and analyses of two other hybrids (1-3 and 1-4) are summarized.



B

```
 1 GCCGCGCTGGTGGCGGCGGCGCGTCGTTGCAGTTGCGCCATCTGTCAGGAGCGGAGCCGGCGAGGAGGGGGCTGCCGCGGGCGAGGAGGAGGGGTCGCCGCGAGCCGAAGGCCTTCGAGA
121 CCCGCCCGCCGCCCGGCGGCGAGAGTAGAGGCGAGGTTGTTGTGCGAGCGGCGCGTCCTCTCCCGCCCGGGCGCGCCGCGCTTCTCCCAGCGCACCGAGGACCGCCCGGGCGCACACAAA
241 GCCGCCGCCCGCGCCGCACCGCCCGGCGGCCGCCGCCCGCGCCAGGGAGGGATTCGGCCGCCGGGCCGGGGACACCCCGGCGCCGCCCCCTCGGTGCTCTCGGAAGGCCCACCGGCTCCC
361 GGGCCCGCCGGGGACCCCCCGGAGCCGCCTCGGCCGCGCCGGAGGAGGGCGGGGAGAGGACCATGTGAGTGGGCTCCGGAGCCTCAGCGCCGCGCAGTTTTTTTGAAGAAGCAGGATGCT

                                                                                                                    1
                                                                                                                    MetGlyCysValGlnCysLys
481 GATCTAAACGTGGAAAAAGACCAGTCCTGCCTCTGTTGTAGAAGACATGTGGTGTATATAAAGTTTGTGATCGTTGGCGGAAATTTTGGAATTTAGATAATGGGCTGTGTGCAATGTAAG

          10                                  20                                  30                                  40
    AspLysGluAlaThrLysLeuThrGluGluArgAspGlySerLeuAsnGlnSerSerGlyTyrArgTyrGlyThrAspProThrProGlnHisTyrProSerPheGlyValThrSerIle
601 GATAAAGAAGCAACAAAACTGACGGAGGAGAGGGACGGCAGCCTGAACCAGAGCTCTGGGTACCGCTATGGCACAGACCCCACCCCTCAGCACTACCCCAGCTTCGGTGTGACCTCCATC

          50                                  60                                  70                                  80
    ProAsnTyrAsnAsnPheHisAlaAlaGlyGlyGlnGlyLeuThrValPheGlyGlyValAsnSerSerSerHisThrGlyThrLeuArgThrArgGlyGlyThrGlyValThrLeuPhe
721 CCCAACTACAACAACTTCCACGCAGCCGGGGGCCAAGGACTCACCGTCTTTGGAGGTGTGAACTCTTCGTCTCATACGGGGACCTTGCGTACGAGAGGAGGAACAGGAGTGACACTCTTT

          90                                  100                                 110                                 120
    ValAlaLeuTyrAspTyrGluAlaArgThrGluAspAspLeuSerPheHisLysGlyGluLysPheGlnIleLeuAsnSerSerGluGlyAspTrpTrpGluAlaArgSerLeuThrThr
841 GTGGCCCTTTATGACTATGAAGCACGGACAGAAGATGACCTGAGTTTTCACAAAGGAGAAAAATTTCAAATATTGAACAGCTCGGAAGGAGATTGGTGGGAAGCCCGCTCCTTGACAACT

          130                                 140                                 150                                 160
    GlyGluThrGlyTyrIleProSerAsnTyrValAlaProValAspSerIleGlnAlaGluGluTrpTyrPheGlyLysLeuGlyArgLysAspAlaGluArgGlnLeuLeuSerPheGly
961 GGAGAGACAGGTTACATTCCCAGCAATTATGTGGCTCCAGTTGACTCTATCCAGGCAGAAGAGTGGTACTTTGGAAAACTTGGCCGAAAAGATGCTGAGCGACAGCTATTGTCCTTTGGA

          170                                 180                                 190                                 200
    AsnProArgGlyThrPheLeuIleArgGluSerGluThrThrLysGlyAlaTyrSerLeuSerIleArgAspTrpAspAspMetLysGlyAspHisValLysHisTyrLysIleArgLys
1081 AACCCAAGAGGTACCTTTCTTATCCGCGAGAGTGAAACCACCAAAGGTGCCTATTCACTTTCTATCCGTGATTGGGATGATATGAAAGGAGACCATGTCAAACATTATAAAATTCGCAAA

          210                                 220                                 230                                 240
    LeuAspAsnGlyGlyTyrTyrIleThrThrArgAlaGlnPheGluThrLeuGlnGlnLeuValGlnHisTyrSerGluArgAlaAlaGlyLeuCysCysArgLeuValValProCysHis
1201 CTTGACAATGGTGGATACTACATTACCACCCGGGCCCAGTTTGAAACACTTCAGCAGCTTGTACAACATTACTCAGAGAGAGCTGCAGGTCTCTGCTGCCGCCTAGTAGTTCCCTGTCAC

          250                                 260                                 270                                 280
    LysGlyMetProArgLeuThrAspLeuSerValLysThrLysAspValTrpGluIleProArgGluSerLeuGlnLeuIleLysArgLeuGlyAsnGlyGlnPheGlyGluValTrpMet
1321 AAAGGGATGCCAAGGCTTACCGATCTGTCTGTCAAAACCAAAGATGTCTGGGAAATCCCTCGAGAATCCCTGCAGTTGATCAAGAGACTGGGAAATGGGCAGTTTGGGGAAGTATGGATG

          290                                 300                                 310                                 320
    GlyThrTrpAsnGlyAsnThrLysValAlaIleLysThrLeuLysProGlyThrMetSerProGluSerPheLeuGluGluAlaGlnIleMetLysLysLeuLysHisAspLysLeuVal
1441 GGTACCTGGAATGGAAACACAAAAGTAGCCATAAAGACTCTTAAACCAGGCACAATGTCCCCCGAATCATTCCTTGAGGAAGCGCAGATCATGAAGAAGCTGAAGCACGACAAGCTGGTC

          330                                 340                                 350                                 360
    GlnLeuTyrAlaValValSerGluGluProIleTyrIleValThrGluTyrMetAsnLysGlySerLeuLeuAspPheLeuLysAspGlyGluGlyArgAlaLeuLysLeuProAsnLeu
1561 CAGCTCTATGCAGTGGTGTCTGAGGAGCCCATCTACATCGTCACCGAGTATATGAACAAAGGAAGTTTACTGGATTTCTTAAAAGATGGAGAAGGAAGAGCTCTGAAATTACCAAATCTT

          370                                 380                                 390                                 400
    ValAspMetAlaAlaGlnValAlaAlaGlyMetAlaTyrIleGluArgMetAsnTyrIleHisArgAspLeuArgSerAlaAsnIleLeuValGlyAsnGlyLeuIleCysLysIleAla
1681 GTGGACATGGCAGCACAGGTGGCTGCAGGAATGGCTTACATCGAGCGCATGAATTATATCCATAGAGATCTGCGATCAGCAAACATTCTAGTGGGGAATGGACTCATATGCAAGATTGCT

          410                                 420                                 430                                 440
    AspPheGlyLeuAlaArgLeuIleGluAspAsnGluTyrThrAlaArgGlnGlyAlaLysPheProIleLysTrpThrAlaProGluAlaAlaLeuTyrGlyArgPheThrIleLysSer
1801 GACTTCGGATTGGCCCGATTGATAGAAGACAATGAGTACACAGCAAGACAAGGTGCAAAGTTCCCCATCAAGTGGACGGCCCCCGAGGCAGCCCTGTACGGGAGGTTCACAATCAAGTCT

          450                                 460                                 470                                 480
    AspValTrpSerPheGlyIleLeuLeuThrGluLeuValThrLysGlyArgValProTyrProGlyMetAsnAsnArgGluValLeuGluGlnValGluArgGlyTyrArgMetProCys
1921 GACGTGTGGTCTTTTGGAATCTTACTCACAGAGCTGGTCACCAAAGGAAGAGTGCCATACCCAGGCATGAACAACCGGGAGGTGCTGGAGCAGGTGGAGCGAGGCTACAGGATGCCCTGC

          490                                 500                                 510                                 520
    ProGlnAspCysProIleSerLeuHisGluLeuMetIleHisCysTrpLysLysAspProGluGluArgProThrPheGluTyrLeuGlnSerPheLeuGluAspTyrPheThrAlaThr
2041 CCGCAGGACTGCCCCATCTCTCTGCATGAGCTCATGATCCACTGCTGGAAAAAGGACCCTGAAGAACGCCCCACTTTTGAGTACTTGCAGAGCTTCCTGGAAGACTACTTTACCGCGACA

          530                 537
    GluProGlnTyrGlnProGlyGluAsnLeuEnd
2161 GAGCCCCAGTACCAACCTGGTGAAAACCTGTAAGGCCCGGGTCTGCGGAGAGAGGCCTTGTCCCAGAGGCTGCCCCACCCCTCCCCATTAGCTTTCAATTCCGTAGCCAGCTGCTCCCCA

2281 GCAGCGGAACCGCCCAGGATCAGATTGCATGTGACTCTGAAGCTGACGAACTTCCATGGCCCTCATTAATGACACTTGTCCCCAAATCCGAACCTCCTCTGTGAAGCATTCGAGACAGAA
2401 CCTTGTTATTTCTCAGACTTTGGAAAATGCATTGTATCGATGTTATGTAAAAGGCCAAACCTCTGTTCAGTGTAAATAGTTACTCCAGTGCCAACAATCCTAGTGCTTTCCTTTTTTAAA
2521 AATGCAAATCCTATGTGATTTTAACTCTGTCTTCACCTGATTCAACTAAAAAAAAAAAGTATTATTTTCCAAAAGTGGCCTCTTTGTCTAAAACAATAAAATTTTTTTTCATGTTTTAAC
2641 AAAAACC
```

FIG. 2. Complete nucleotide and amino acid sequences of c-*syn*. (*A*) Four overlapping cDNA clones of c-*syn*. The open box indicates the predicted coding frame. (*B*) Nucleotide sequence of λSN-2 insert and its deduced amino acid sequence. Nucleotides are numbered at the left. Amino acids are numbered starting with the initiation methionine. Termination codons, in the 5′-untranslated region are underlined. Besides the translated sequence, two fairly long open reading frames were identified in the 5′-untranslated region, one with 192 codons and one with 142 codons, which are located between positions 1 and 576 and between 3 and 428, respectively, in the nucleotide sequence. No initiator methionine was available for these open frames. The AATAAA sequence at the 3′ end is underlined.

of the c-syn sequence with that of chicken c-src (23), human c-fgr (9), and v-yes (31). The homology is prominent in the amino acid sequence between position 83 and 537 of syn: 80% (syn/yes), 77% (syn/src), and 77% (syn/fgr). Since this domain of *src* as well as of *yes* and *fgr* has been shown to encode protein with protein-tyrosine kinase activity, the syn protein is also expected to specify the kinase activity. This conclusion is supported by the fact that a lysine residue at

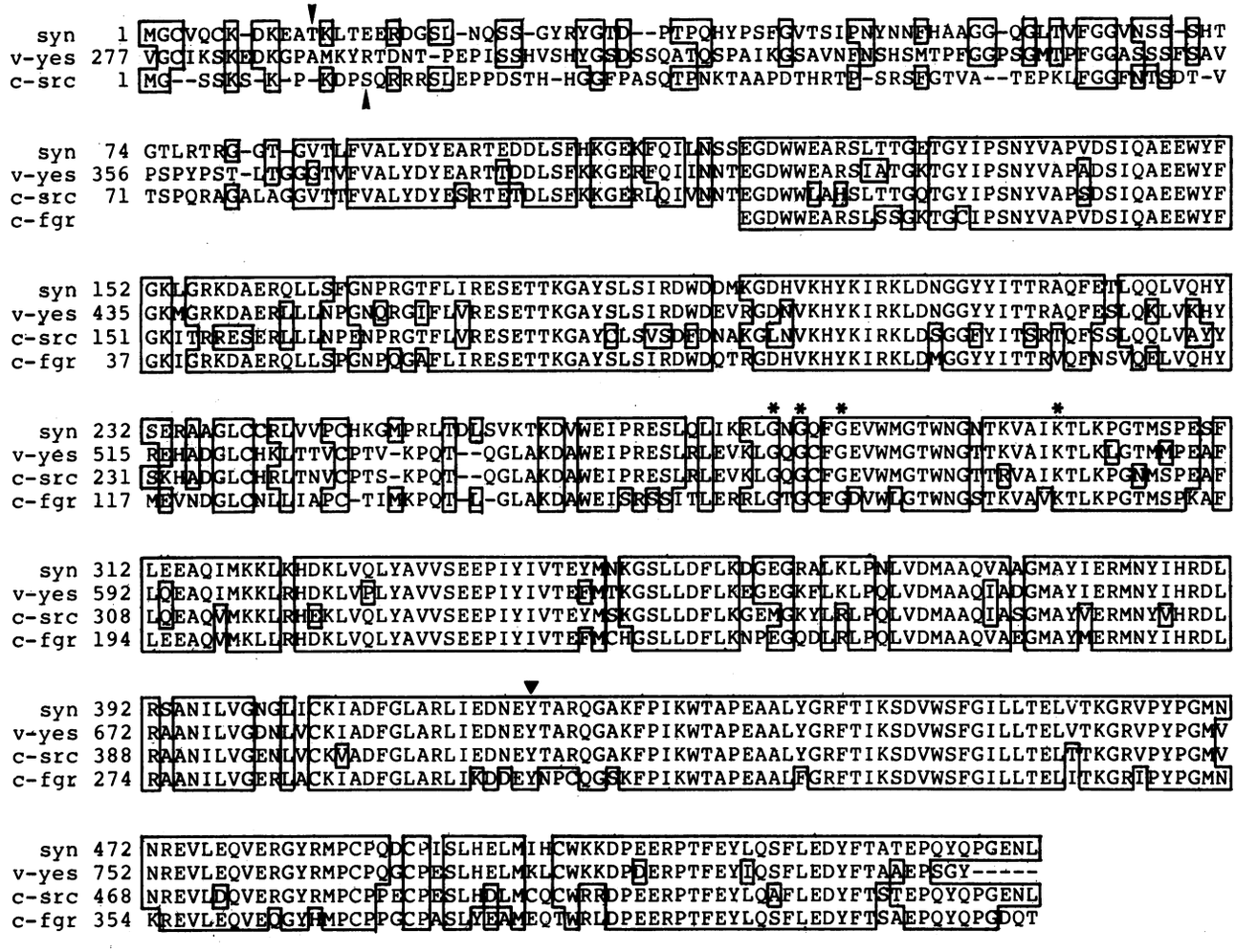


FIG. 3. Comparison of the amino acid sequences of human c-syn, chicken c-src (23), v-yes (31), and human c-fgr (9). Identical amino acids in the putative c-syn protein and other proteins are boxed. Asterisks indicate residues that are thought to be associated with nucleotide binding. The closed triangle indicates a tyrosine expected to be autophosphorylated. Ser-12 of $pp60^{src}$, which is phosphorylated by protein kinase C, and the corresponding amino acid residue, Thr-12, of c-syn are indicated by arrows. Amino acids of the four proteins are numbered on the left. The extreme 5′ sequence of the human c-fgr protein is not available.

position 299, suggested to be involved in ATP binding (32), is conserved in the syn protein sequence, as is a consensus sequence Gly-Xaa-Gly-Xaa-Xaa-Gly (33) at position 278–283 (Fig. 3). Furthermore, the Tyr-416 autophosphorylation site of $pp60^{src}$ (34) is also conserved in the putative syn protein (tyrosine at position 420). The amino acid homology between syn and yes in the amino-terminal moiety (positions 1–82) is extremely low (29%). Homology between syn and src is also low (23%) at this region. The low homology of the amino-terminal portion of the syn protein to those of related proteins suggests that the amino-terminal sequences specify the specialized functions of these proteins. The amino acid at position 81 of chicken c-src (equivalent to position 82 of c-syn) precisely coincides with the splicing site in the chicken c-*src* gene (23). The splicing mechanism could have been involved in a gene-shuffling process that may have played a role in the generation of *src*/*yes*-related genes with distinct functions. Threonine at position 12 is flanked by lysine residues and may well be phosphorylated by protein kinase C as reported for Ser-12 of $pp60^{src}$ (35). However, it should be mentioned that Thr-12 of syn is embedded in the more acidic environment than is Ser-12 of the latter.

Expression of the c-*syn* Gene. Poly(A)$^+$ RNA was prepared from human placenta, adenocarcinoma cells of the human pancreas UCVA, human embryo fibroblasts, chronic myelogenous leukemia cells K562, human lymphoid cells IM-9, epidermoid carcinoma cells A431, and human follicular lymphoma cells FL18. Blot-hybridization analysis of these RNAs with the KD probe revealed the expression of the 2.8-kb syn mRNA at various levels in these cells, although it was not detected in UCVA cells (Fig. 4*A*). The size of the syn transcript differs from those of the c-yes (7), c-src (36), and c-fgr (9) mRNAs. In the same filter, the c-fgr mRNA was detected exclusively in IM-9 (Fig. 4*C*). Only a low level of the c-fgr transcript was observed in placenta (9). As IM-9 cells have been immortalized with the infection by Epstein–Barr virus, our observation is consistent with the finding that c-fgr mRNA is induced by Epstein–Barr virus infection (37). Transcription of the c-*yes* gene was relatively high in both A431 and UCVA cells, in which the expression of the epidermal growth receptor at the cell surface is high and was undetectable in FL18 cells. (Fig. 4*B*) (ref. 7). Comparison of the hybridization blots of these mRNAs suggests that the expression of these related genes is distinctly regulated in each of the cells examined. Relatively high expression of the c-syn and c-fgr mRNAs in IM9 deserves further analysis.

Implications. In this report, we showed identification of a *yes*-related gene that could encode a protein-tyrosine kinase. Besides *syn*, we have recently obtained cDNA clones carrying distinct *yes*-related sequences from chicken and human cDNA libraries (unpublished data). These genes also can encode proteins with protein-tyrosine kinase activity. In addition, a lymphocyte-specific protein-tyrosine kinase gene *lsk*^{T}/*tck* recently has been cloned molecularly and shown to

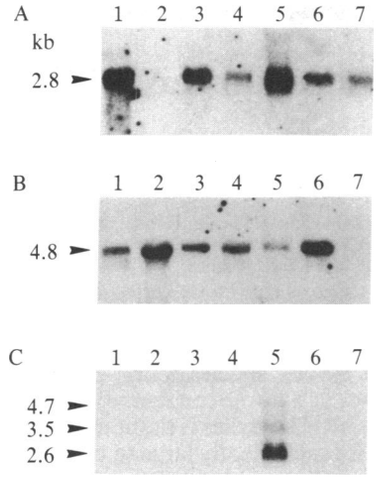


FIG. 4. Expression of the c-*syn* gene in human cells. Poly(A)$^+$ RNAs (5 μg) were subjected to blot hybridization with ^{32}P-labeled KD DNA probe (*A*), c-*yes* probe (unpublished data) (*B*), and c-*fgr* probe (9) (*C*) under stringent conditions as described (15). RNAs were isolated from human placenta (lane 1), UCVA (lane 2), human embryo fibroblast (lane 3), K562 (lane 4), IM-9 (lane 5), A431 (lane 6), and FL18 (lane 7) cells. ^{32}P-labeled fragments from a *Hin*dIII digest of phage λ DNA were used as size markers.

be highly related to the *yes* gene (38, 39). We do not know how the expression of these "similar but distinct" genes is regulated. Protein-tyrosine kinases are rare enzymes and have been believed to have a regulatory function in normal cells. An idea that stems from the findings of higher expression of the c-*src* gene in neural tissues (40, 41) is that the "nonreceptor"-type protooncogene of the kinase family may be involved in a variety of functions specific to differentiated cells. As far as we examined, expression of the c-*fgr* gene was specific in IM9 human lymphoid cells. In contrast, c-*yes* and c-*syn* expression of mRNA was observed in various types of cells, although the modes of expression of these genes differ from each other. More extensive analysis of these similar but distinct genes, including *in situ* analyses of mRNAs and proteins, is required to determine whether the protein-tyrosine kinase activities of these products are specific to certain cell lineage or whether it plays a more general role in cell growth.

We thank S. Nagata for valuable discussions and S. Sasaki, M. Ando, and M. Ide for typing this manuscript. This work was supported in part by a Grant-in-Aid for Special Project Research Cancer-Bioscience from the Ministry of Education, Science and Culture of Japan.

1. Hunter, T. & Copper, J. A. (1985) *Annu. Rev. Biochem.* **54,** 897–930.
2. Cohen, S., Carpenter, G. & King, L. (1980) *J. Biol. Chem.* **255,** 4834–4842.
3. Kasuga, M., Zick, Y., Blithe, D. L., Crettaz, M. & Kahn, C. R. (1982) *Nature (London)* **298,** 667–669.
4. Nishimura, J., Hung, J. S. & Deuel, T. F. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 4303–4307.
5. Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D. & Seeburg, P. H. (1984) *Nature (London)* **309,** 418–425.
6. Sherr, C. J., Rettenmier, C. W., Sacca, R., Roussel, M. F., Look, A. T. & Stanley, E. R. (1985) *Cell* **41,** 665–676.
7. Semba, K., Yamanashi, Y., Nishizawa, M., Sukegawa, J., Yoshida, M. C., Sasaki, M., Yamamoto, T. & Toyoshima, K. (1985) *Science* **227,** 1038–1040.
8. Yoshida, M. C., Sasaki, M., Mise, K., Semba, K., Nishizawa, M., Yamamoto, T. & Toyoshima, K. (1985) *Jpn. J. Cancer Res. (Gann)* **76,** 559–562.
9. Nishizawa, M., Semba, K., Yoshida, M. C., Yamamoto, T., Sasaki, M. & Toyoshima, K. (1986) *Mol. Cell. Biol.* **6,** 511–517.
10. Tronick, S. R., Popescu, N. C., Cheah, M. S. C., Swan, D. C., Amsbaugh, S. C., Lengel, C. R., DiPaolo, J. A. & Robbins, K. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6595–6599.
11. Gamou, S., Kim, Y. S. & Shimizu, N. (1984) *Mol. Cell. Endocrinol.* **37,** 205–213.
12. Ohno, H., Doi, S., Fukuhara, S., Nishikori, M., Uchino, H. & Fujii, H. (1985) *Jpn. J. Cancer Res. (Gann)* **76,** 563–566.
13. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connel, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15,** 687–701.
14. Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA Cloning*, ed. Glover, D. M. (IRL, Oxford), Vol. 1, pp. 49–78.
15. Semba, K., Kamata, N., Toyoshima, K. & Yamamoto, T. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6497–6501.
16. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
17. Mizusawa, S., Nishimura, S. & Seela, F. (1986) *Nucleic Acids Res.* **14,** 1319–1324.
18. Hattori, M. & Sakaki, Y. (1986) *Anal. Biochem.* **152,** 232–238.
19. Davidson, R. L. & Gerald, P. S. (1976) *Somat. Cell Genet.* **2,** 165–176.
20. Yoshida, M. C., Ikeuchi, T. & Sasaki, M. (1975) *Proc. Jpn. Acad.* **51,** 184–187.
21. Gibbs, C. P., Tanaka, A., Anderson, S. K., Radul, J., Baar, J., Ridgway, A., Kung, H.-J. & Fujita, D. J. (1985) *J. Virol.* **53,** 19–24.
22. Mount, S. M. (1982) *Nucleic Acids Res.* **10,** 459–472.
23. Takeya, T. & Hanafusa, H. (1983) *Cell* **32,** 881–890.
24. Anderson, S. K., Gibbs, C. P., Tanaka, A., Kung, H.-J. & Fujita, D. J. (1985) *Mol. Cell. Biol.* **5,** 1122–1129.
25. Wang, J. Y. J., Ledley, F., Goff, S., Lee, R., Groner, Y. & Baltimore, D. (1984) *Cell* **36,** 349–356.
26. Huang, C.-C., Hammond, C. & Bishop, J. M. (1985) *J. Mol. Biol.* **181,** 175–186.
27. Bonner, T. I., Kerby, S. B., Sutrave, P., Gunnel, M. A., Mark, G. & Rapp, U. R. (1985) *Mol. Cell. Biol.* **5,** 1400–1407.
28. Le Beau, M. M., Westbrook, C. A., Diaz, M. O. & Rowley, J. D. (1984) *Nature (London)* **312,** 70–71.
29. Yunis, J. J. (1983) *Science* **221,** 227–236.
30. Kozak, M. (1984) *Nucleic Acids Res.* **12,** 857–872.
31. Kitamura, N., Kitamura, A., Toyoshima, K., Hirayama, Y. & Yoshida, M. (1982) *Nature (London)* **297,** 205–208.
32. Kamps, M. P., Taylor, S. S. & Sefton, B. M. (1984) *Nature (London)* **310,** 589–592.
33. Wierenga, R. K. & Hol, W. G. J. (1983) *Nature (London)* **302,** 842–844.
34. Smart, J. E., Oppermann, H., Czernilofsky, A. P., Purchio, A. F., Erikson, R. L. & Bishop, J. M. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6013–6017.
35. Gould, K. L., Woodgett, J. R., Cooper, J. A., Buss, J. E., Shalloway, D. & Hunter, T. (1985) *Cell* **42,** 849–857.
36. Parker, R. C., Mardon, G., Lebo, R. V., Varmus, H. E. & Bishop, J. M. (1985) *Mol. Cell. Biol.* **5,** 831–838.
37. Cheah, M. S. C., Ley, T. J., Tronick, S. R. & Robbins, K. C. (1986) *Nature (London)* **319,** 238–240.
38. Marth, J. D., Peet, R., Krebs, E. G. & Perlmutter, R. M. (1985) *Cell* **43,** 393–404.
39. Voronova, A. F. & Sefton, B. M. (1986) *Nature (London)* **319,** 682–685.
40. Brugge, J. S., Cotton, P. C., Queral, A. E., Barrett, J. N., Nonner, D. & Keane, R. W. (1985) *Nature (London)* **316,** 554–557.
41. Simon, M. A., Drees, B. Kornberg, T. & Bishop, J. M. (1985) *Cell* **42,** 831–840.